

The complexity of the virus world

Eugene V. Koonin, Yuri I. Wolf, Keizo Nagasaki and Valerian V. Dolja

In their interesting correspondence on our Analysis article (The Big Bang of picorna-like virus evolution antedates the radiation of eukaryotic supergroups. *Nature Rev. Microbiol.* **6**, 925–939 (2008))¹, Krupović and Bamford insist that any study on the evolution and classification of viruses should take a capsid-centric approach (Does the evolution of viral polymerases reflect the origin and evolution of viruses? *Nature Rev. Microbiol.* 9 Feb 2008 (doi:10.1038/nrmicro2030-c1))². From this perspective, they question the relevance of our recent study of the picorna-like virus superfamily, which was based on an analysis of the evolution of proteins involved in viral genome replication¹. The importance of the capsid in the life cycle of viruses is obvious, and the recurrence of the capsid protein fold (the jelly-roll domain) in an enormous range of icosahedral viruses is remarkable^{3,4}. Having said this, we think that the capsid-centric approach and the associated definition of a virus⁴ are unnecessarily rigid and could lead to neglect of important aspects of the evolutionary dynamics of the virus world (virosphere) and even to overt contradictions. A more inclusive approach to the evolution of viruses that considers genes with different functions is likely to be more productive.

In an earlier study of the virus world, we delineated a set of viral hallmark genes that included both the jelly-roll capsid protein and the RNA-dependent RNA polymerases (RdRps), as well as several additional genes, such as those that encode superfamily 3 helicases and packaging ATPases⁵. The hallmark genes are present in extremely diverse groups of viruses, in which they are central to viral replication and structure, whereas cellular life forms possess only distant homologues of these genes. Differential mixing and matching of the hallmark genes and their combination with other, lineage-specific genes provides for the enormous diversity of genome organizations of viruses and virus-like elements. Therefore, none of the hallmark genes alone, be it the capsid protein or the RdRp, represents the ‘essence’ of a virus; less metaphorically, the phylogeny of any single gene cannot adequately represent the evolutionary processes in the virus world. These processes constitute a complex

network of vertical and horizontal paths that, in principle, cannot be described by a single tree, which represents a paradigm shift that is increasingly recognized in evolutionary biology in general^{6,7}.

In our reconstruction of the evolution of the picorna-like superfamily of RNA viruses, which was defined by the analysis of overlapping arrays of conserved genes, rather than a single gene, we attempted to infer the history of vertical inheritance and horizontal gene transfers¹. Although the phylogenies of even the most highly conserved genes cannot, in principle, fully reflect the evolutionary process, some provide an indispensable framework on which to map various events that occur during the evolution of a particular class of viruses. From this perspective, for RNA viruses, the RdRp seems to have a distinct advantage over the capsid protein. Indeed, except for some satellite viruses, all RNA viruses encode the RdRp, because cellular life forms possess no enzymes that are capable of replicating long RNA molecules. By contrast, RNA elements, which encode no capsid protein but clearly evolved from RNA viruses, are common, as are displacements of the capsid protein. The most conspicuous and relevant cases are the capsid-less hypoviruses and the potyviruses, which possess filamentous capsids. The inclusion of potyviruses in the picorna-like superfamily is hard to question given not only the high conservation of the RdRp sequence but also the presence of two other conserved genes, the 3C-like protease and VPg. The hypoviruses have a similar genome organization to, and in all likelihood evolved from, potyviruses: the capsid protein was probably lost during the adaptation to fungi⁸. In the case of the hypoviruses, the vector of evolution is clear, and the loss of the capsid protein gene makes the capsid-centric perspective irrelevant. In addition to the hypoviruses, there are several groups of capsid-less RNA elements that seem to be derived from other viruses, including narnaviruses, mitoviruses and umbraviruses; all these groups are recognized as viruses by the International Committee on Taxonomy of Viruses⁹. Extensive connections between bona fide viruses and capsid-less elements are also

observed among retroird viruses and small DNA viruses⁴.

The statement that “the jelly-roll fold is common to most ssRNA eukaryotic viruses” is inaccurate. Even apart from the negative-strand RNA viruses, none of which is known to possess this fold, most of the positive-strand RNA viruses within the order Nidovirales and the alphavirus-like superfamily have capsid or nucleocapsid proteins other than the jelly-roll^{10,11}. The alphavirus-like superfamily is defined by a conserved core of replicative genes that encode the capping enzyme, superfamily 1 helicase and RdRp. The animal viruses of this superfamily (alphaviruses) possess capsid proteins that belong to the family of chymotrypsin-like proteases, whereas most of the plant viruses, such as tobamoviruses, flexiviruses and closteroviruses, form helical capsids; none of these capsid proteins is related to the jelly-roll. Thus, owing to horizontal gene transfer and gene displacement, capsids show greater variance in distribution than RdRp and other RNA replication enzymes that, accordingly, are more useful for parsimonious reconstructions of RNA virus evolution.

To conclude, the virus world is a dynamic network of relationships in which genes have diverse, variously intertwined histories and viruses that possess capsids and mobile genetic elements that do not show a recurring pattern of evolutionary links. The capsid-centric approach is therefore unable to capture the full complexity of the virosphere evolution and the contributions it makes to the evolution of cellular life forms.

Eugene V. Koonin and Yuri I. Wolf are at the National Center for Biotechnology Information, National Institutes of Health, Bethesda, Maryland 20894, USA.

Keizo Nagasaki is at the National Research Institute of Fisheries and Environment of Inland Sea, 2-17-5 Maruishi, Hiroshima, 739-0452, Japan.

Valerian V. Dolja is at the Department of Botany and Plant Pathology and Center for Genome Research and Biocomputing, Oregon State University, Corvallis, Oregon 97331, USA.

Correspondence to V.V.D. and E.V.K. e-mails: doljav@science.oregonstate.edu; koonin@ncbi.nlm.nih.gov

1. Koonin, E. V., Wolf, Y. I., Nagasaki, K. & Dolja, V. V. The Big Bang of picorna-like virus evolution antedates the radiation of eukaryotic supergroups. *Nature Rev. Microbiol.* **6**, 925–939 (2008).
2. Krupović, M. & Bamford, D. H. Does the evolution of viral polymerases reflect the origin and evolution of viruses? *Nature Rev. Microbiol.* 9 Feb 2008 (doi:10.1038/nrmicro2030-c1).
3. Krupović, M. & Bamford, D. H. Virus evolution: how far does the double β-barrel viral lineage extend? *Nature Rev. Microbiol.* **6**, 941–948 (2008).

CORRESPONDENCE

4. Raoult, D. & Forterre, P. Redefining viruses: lessons from Mimivirus. *Nature Rev. Microbiol.* **6**, 315–319 (2008).
5. Koonin, E. V., Senkevich, T. G. & Dolja, V. V. The ancient virus world and evolution of cells. *Biol. Direct* **1**, 29 (2006).
6. Doolittle, W. F. & Baptiste, E. Pattern pluralism and the Tree of Life hypothesis. *Proc. Natl Acad. Sci. USA* **104**, 2043–2049 (2007).
7. Koonin, E. V. & Wolf, Y. I. Genomics of bacteria and archaea: the emerging dynamic view of the prokaryotic world. *Nucleic Acids Res.* **36**, 6688–6719 (2008).
8. Koonin, E. V., Choi, G. H., Nuss, D. L., Shapira, R. & Carrington, J. C. Evidence for common ancestry of a chestnut blight hypovirulence-associated double-stranded RNA and a group of positive-strand RNA plant viruses. *Proc. Natl Acad. Sci. USA* **88**, 10647–10651 (1991).
9. Fauquet, C. M., Mayo, M. A., Maniloff, J., Desselberger, U. & Ball, L. A. Virus Taxonomy. Eighth Report of the International Committee on Taxonomy of Viruses (Elsevier, Amsterdam, 2005).
10. Koonin, E. V. & Dolja, V. V. Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. *Crit. Rev. Biochem. Mol. Biol.* **28**, 375–430 (1993).
11. Gorbalenya, A. E., Enjuanes, L., Ziebuhr, J. & Snijder, E. J. Nidovirales: evolving the largest RNA virus genome. *Virus Res.* **117**, 17–37 (2006).